ABSTRACT

The present invention relates to a system and method to facilitate data mining applications and automated evaluation of models for continuous variable data. In one aspect, a system is provided that facilitates decision tree learning. The system includes a learning component that generates non-standardized data that relates to a split in a decision tree and a scoring component that scores the split as if the non-standardized data at a subset of leaves of the decision tree had been shifted and/or scaled. A modification component can also be provided for a respective candidate split score on the decision tree, wherein the above data or data subset can be modified by shifting and/or scaling the data and a new score is computed on the modified data. Furthermore, an optimization component can be provided that analyzes the data and determines whether to treat the data as if it was: (1) shifted, (2) scaled, or (3) shifted and scaled.